

Data  $\Rightarrow$  Facts

Information  $\Rightarrow$  Meaningful Facts

Knowledge  $\Rightarrow$  Application of meaningful Facts

Wisdom  $\Rightarrow$  make right decision

B.I.  $\Rightarrow$  Business Intelligence  $\Rightarrow$  data mining

OLAP is not data mining

Data Mining  $\Rightarrow$  Search for (predict) future data behavior (pattern) from the past data pattern, and come up with a (business) rule.

such as If A event occurs  
then B event occurs

Coding  $\Rightarrow$  digitalize data for data mining  
or programming

Enrichment  $\Rightarrow$  Find out related information of data

Candidate generation  $\Rightarrow$  potential data will be used in a Rule (Relevancy)

Support level  $\Rightarrow$  Percentage of relevant data out of total data

Confidence level  $\Rightarrow$  Accuracy of prediction of two events occur together given one event already occurs

Apriori algorithm  $\Rightarrow$  Given confidence & support levels

What is the Association Rule of the source data by computer automation.

Starts with single event, and increment event by one in each loop until the event cannot meet the support and confidence level.

Apriori algorithm starts with non-candidate data set.

Frequent Pattern Growth tree starts with candidate data set.